DIALib-QC v1.2 Tutorial

Here we present a spectral ion library (*K562_q3bad_PV*) referring from the Midha *et. al.* paper published in Nature communication (2020).

CITE DIALib-QC

DIALib-QC-Github: DIALib-QC github

DIALib-QC-Discuss at Google Groups: DIALib-QC Google discussion group

A) All the below mentioned steps works well with **PeakView, Spectronaut and OpenSWATH library** formats. Please download all the files to test running DIALib-QC.

Step 1: Objective: Assessment of the ion library (*K562_q3bad_PV.txt*)

perl assess_swathlib.pl --ion_library K562_q3bad_PV.txt --full_stats --rt_stats --assess_massdiff --coldefs --invert_output --swath_file SWATH_100.txt --peptide_file Uniprot-Human.ppeps

DIALib-QC flagged fragment ions average mass error of .0064 in the *K562_q3bad_PV.txt.QC* report.

library	K562_q3bad_PV.txt	Name of library file being analyzed				
precursor_ok	36431	Number of assays where precursor is within 5 PPM (parts per million m/z) of theoretical				
precursor_bad	0	Number of assays where fragment is more than 5 PPM from theoretical				
fragment_ok	9837	Number of assays where fragment is within 1 PPM of theoretical				
fragment_bad	208749	Number of assays where precursor is more than 1 PPM from theoretical				
fragment_na	0	Number of assays where peak annotation not found in expected b/y series				
fragment_avg_mdiff	0.0064	Average m/z difference between reported and theoretical fragment				
swa_defined	36431	Number of peptide ions that fall into a defined SWATH bin				
swa_missing	0	Number of peptide ions that fail to fall into a defined SWATH bin (Pepions - swa_defined)				
swa_conflict	0	Number of fragment_ions that fall into same SWATH(s) as precursor				
swa_ok	255528	Number of fragment_ions that do not fall into same SWATH(s) as precursor				
swa_conflict_assay	0	Number of precursor that have at least one failing fragment				
swa_5	655	Number of fragment ions that fall within 5 Th of precursor ion				
swa_25	9397	Number of fragment ions that fall within 25 Th of precursor ion				
problem_assays	36431	Assays whose precursor or any fragment m/z values do not match SWATHs file or differ significantly from theoretical values				

Step 2: Objective: Correct fragment (q3) mass error in the ion library using theoretical masses.

perl assess_swathlib.pl --ion_library K562_q3bad_PV.txt --assess_massdiff --correct_mz --swath_file SWATH_100.txt --peptide_file Uniprot-Human.ppeps

We assess the *K562_q3bad_PV.txt.mz_corrected* new library (output from this step)

perl assess_swathlib.pl --ion_library K562_q3bad_PV.txt.mz_corrected --full_stats --rt_stats --assess_massdiff --coldefs --invert_output --swath_file SWATH_100.txt --peptide_file Uniprot-Human.ppeps

Now there is no fragment ion mass error but DIALib-QC flagged few conflict assays in the *K562_q3bad_PV.txt.mz_corrected.QC* report.

library	K562_q3bad_PV.txt.mz_corrected	Name of library file being analyzed
precursor_ok	36431	Number of assays where precursor is within 5 PPM (parts per million m/z) of theoretical
precursor_bad	0	Number of assays where fragment is more than 5 PPM from theoretical
fragment_ok	218586	Number of assays where fragment is within 1 PPM of theoretical
fragment_bad	0	Number of assays where precursor is more than 1 PPM from theoretical
fragment_na	0	Number of assays where peak annotation not found in expected b/y series
fragment_avg_mdiff	0	Average m/z difference between reported and theoretical fragment
swa_defined	36431	Number of peptide ions that fall into a defined SWATH bin
swa_missing	0	Number of peptide ions that fail to fall into a defined SWATH bin (Pepions - swa_defined)
swa_conflict	5	Number of fragment_ions that fall into same SWATH(s) as precursor
swa_ok	255523	Number of fragment_ions that do not fall into same SWATH(s) as precursor
swa_conflict_assay	5	Number of precursor that have at least one failing fragment
swa_5	652	Number of fragment ions that fall within 5 Th of precursor ion
swa_25	9395	Number of fragment ions that fall within 25 Th of precursor ion
problem_assays	5	Assays whose precursor or any fragment m/z values do not match SWATHs file or differ significantly from theoretical values

Step 3: Objective: Remove conflict assays from the library

perl assess_swathlib.pl --ion_library K562_q3bad_PV.txt.mz_corrected --assess_massdiff --filter_assays --swath_file SWATH_100.txt --peptide_file Uniprot-Human.ppeps

We assess the *K562_q3bad_PV.txt.mz_corrected.clean* new library (output from this step)

perl assess_swathlib.pl --ion_library K562_q3bad_PV.txt.mz_corrected.clean --full_stats --rt_stats --assess_massdiff --coldefs --invert_output --swath_file SWATH_100.txt --peptide_file Uniprot-Human.ppeps

Now there is no fragment ion mass error or conflict assays in the *K562_q3bad_PV.txt.mz_corrected.QC* report.

library	K562_q3bad_PV.txt.mz_corrected.clean	Name of library file being analyzed
precursor_ok	36426	Number of assays where precursor is within 5 PPM (parts per million m/z) of theoretical
precursor_bad	0	Number of assays where fragment is more than 5 PPM from theoretical
fragment_ok	218556	Number of assays where fragment is within 1 PPM of theoretical
fragment_bad	0	Number of assays where precursor is more than 1 PPM from theoretical
fragment_na	0	Number of assays where peak annotation not found in expected b/y series
fragment_avg_mdiff	0	Average m/z difference between reported and theoretical fragment
swa_defined	36426	Number of peptide ions that fall into a defined SWATH bin
swa_missing	0	Number of peptide ions that fail to fall into a defined SWATH bin (Pepions - swa_defined)
swa_conflict	0	Number of fragment_ions that fall into same SWATH(s) as precursor
swa_ok	255498	Number of fragment_ions that do not fall into same SWATH(s) as precursor
swa_conflict_assay	0	Number of precursor that have at least one failing fragment
swa_5	649	Number of fragment ions that fall within 5 Th of precursor ion
swa_25	9393	Number of fragment ions that fall within 25 Th of precursor ion
problem_assays	0	Assays whose precursor or any fragment m/z values do not match SWATHs file or differ significantly from theoretical values

Step 4: Objective: To generate graphical plots (*K562_q3bad_PV.txt.mz_corrected.clean.pdf*) summarizing library characteristics

perl DIALib-QC_RPlot.pl K562_q3bad_PV.txt.mz_corrected.clean.fullstats K562_q3bad_.txt.mz_corrected.clean.RT
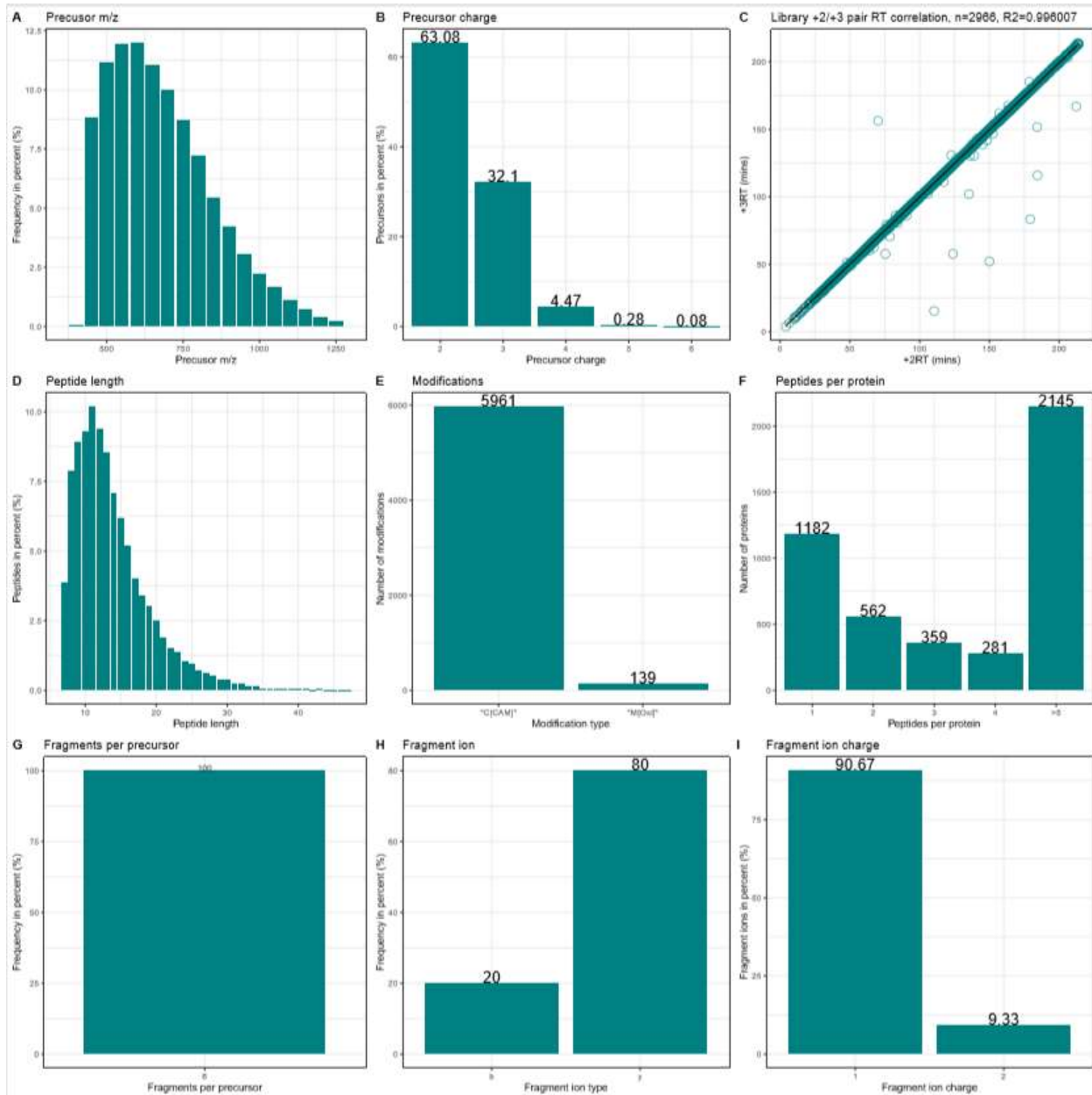
B) In-silico predicted library in Prosit (generic text: Spectronaut compatible)

We used a larger spectral ion library *HeLa_ce28.prosit.library.csv* submitted in iProx database (Chen *et. al.* 2019).

Step 1: Objective: Assessment of the ion library (*HeLa_ce28.prosit.library.csv*)

```
perl assess_swathlib.pl --assess_massdiff  --full_stats --rt_stats --coldefs --invert_output --ion_library
HeLa_ce28.prosit.library.csv
```

DIALib-QC flagged few problem assays in the *HeLa_ce28.prosit.library.csv.QC* report because of difference with theoretical masses.

library	HeLa_ce28.prosit.library
precursor_ok	83976
precursor_bad	0
fragment_ok	5038047
fragment_bad	8243
fragment_na	0
fragment_avg_mdiff	0
problem_assays	7742

Step 2: Objective: Correct to theoretical masses.

perl assess_swathlib.pl --ion_library HeLa_ce28.prosit.library.csv --assess_massdiff --correct_mz

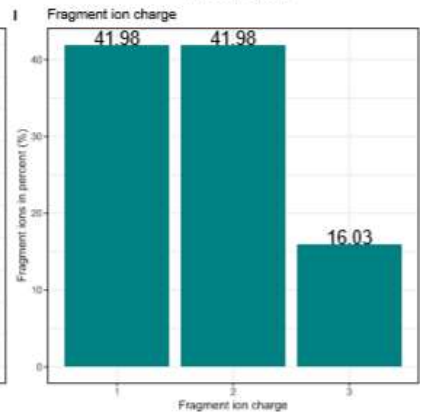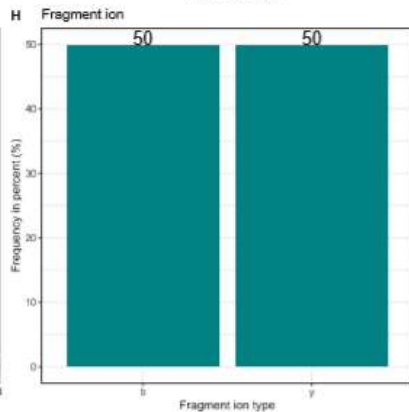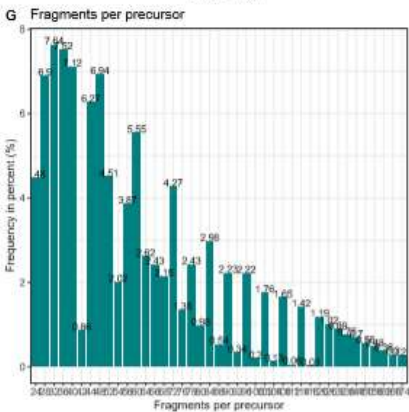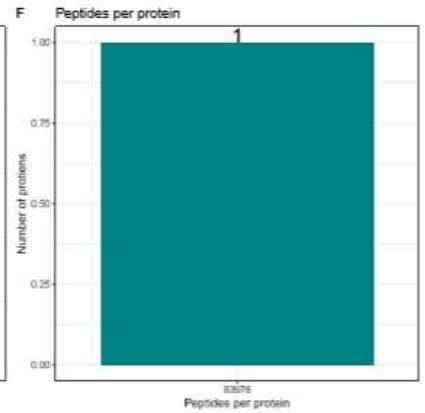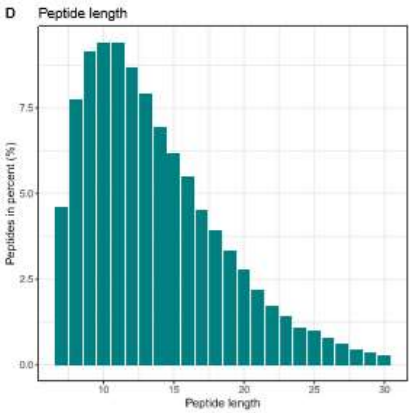We assess the new mass corrected *HeLa_ce28.prosit.library.csv.mz_corrected* library (output from this step)
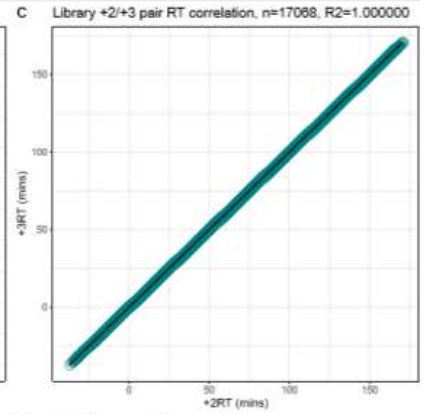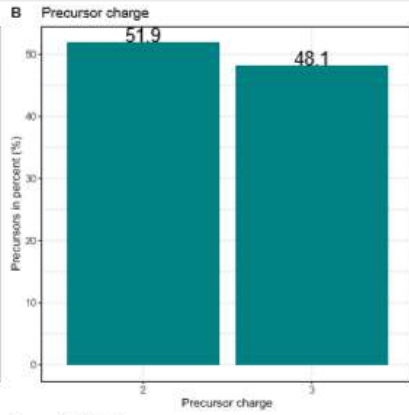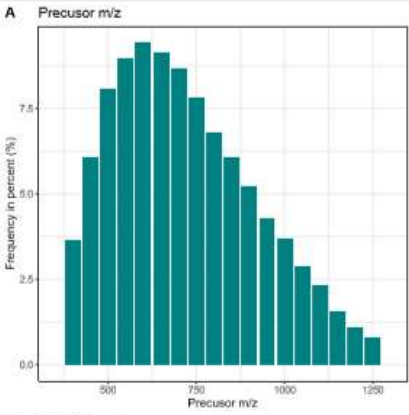
perl assess_swathlib.pl --ion_library HeLa_ce28.prosit.library.csv.mz_corrected --full_stats --rt_stats --assess_massdiff --coldefs --invert_output

Now there is no fragment ion mass error and no problem assays in the mass corrected *HeLa_ce28.prosit.library.csv.mz_corrected.QC* report.

library	HeLa_ce28.prosit.library.csv.mz_corrected		
precursor_ok	83976		
precursor_bad	0		
fragment_ok	5046290		
fragment_bad	0		
fragment_na	0		
fragment_avg_mdiff	0		
problem_assays	0		

Step 3: Objective: To generate graphical plots (*HeLa_ce28.prosit.library.csv.mz_corrected.pdf*) summarizing *in-silico* predicted ion library (free form any error) characteristics

perl DIALib-QC_RPlot.pl HeLa_ce28.prosit.library.csv.mz_corrected.fullstats HeLa_ce28.prosit.library.csv.mz_corrected.RT

A Precursor m/z

B Precursor charge

C Library +2/+3 pair RT correlation, n=17068, R2=1.000000

D Peptide length

E Modifications

F Peptides per protein

G Fragments per precursor

H Fragment ion

I Fragment ion charge
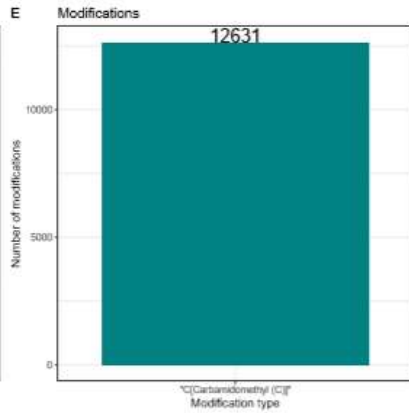
Addition of user defined modification into DIALib-QC version 1.2 code

DIALib-QC perl script (assess_swathlib.pl) contains a comment line #146 which can be changed by the user to incorporate user defined modification.

The information required to replace this line is:

- *Modification Annotation* like *'C[CAM]', 'M[Oxi]', Y[Dhy]* etc;
- *Monoisotopic mass* referred to as 'mz' and
- *Single letter code* for every amino acid '*aa*'.

Please use 'x' as associated 'aa' value for the terminal modifications like *'[+1R]-' [1Ac]-'* etc .

Here is the example

```
# 'Y[Dhy]' => {'mz' => -18.010565, 'aa' => 'y'},#EXAMPLE LINE TO ADD MODIFICATION
```

Desired modification to add is *'H[CAM]'* with monoisotopic mass of *57.0215*

So the above line will be modified to as follows:

```
'H[CAM]'  =>  {'mz' => 57.0215,  'aa' => 'H'}#User defined modification
```

Let us know if you have any queries related to modification addition at DIALib-QC Google discussion group